



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/820,714A

DATE: 10/06/2004

TIME: 12:22:57

Input Set : A:\251697.ST25.txt

Output Set: N:\CRF4\10062004\J820714A.raw

3 <110> APPLICANT: KAO CORPORATION  
 4 Okuda, Mitsuyoshi  
 5 Kobayashi, Tohru  
 6 Sumitomo, Nobuyuki  
 7 Takimura, Yasushi  
 8 Sato, Tsuyoshi  
 10 <120> TITLE OF INVENTION: ALKALINE PROTEASE  
 12 <130> FILE REFERENCE: 251697US0  
 14 <140> CURRENT APPLICATION NUMBER: 10/820,714A  
 15 <141> CURRENT FILING DATE: 2004-04-09  
 17 <150> PRIOR APPLICATION NUMBER: 2003-106709  
 18 <151> PRIOR FILING DATE: 2003-04-10  
 20 <160> NUMBER OF SEQ ID NOS: 24  
 22 <170> SOFTWARE: PatentIn version 3.2  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 434  
 26 <212> TYPE: PRT  
 27 <213> ORGANISM: Bacillus sp. KSM-KP43  
 29 <400> SEQUENCE: 1  
 31 Asn Asp Val Ala Arg Gly Ile Val Lys Ala Asp Val Ala Gln Ser Ser  
 32 1 5 10 15  
 35 Tyr Gly Leu Tyr Gly Gln Gly Gln Ile Val Ala Val Ala Asp Thr Gly  
 36 20 25 30  
 39 Leu Asp Thr Gly Arg Asn Asp Ser Ser Met His Glu Ala Phe Arg Gly  
 40 35 40 45  
 43 Lys Ile Thr Ala Leu Tyr Ala Leu Gly Arg Thr Asn Asn Ala Asn Asp  
 44 50 55 60  
 47 Thr Asn Gly His Gly Thr His Val Ala Gly Ser Val Leu Gly Asn Gly  
 48 65 70 75 80  
 51 Ser Thr Asn Lys Gly Met Ala Pro Gln Ala Asn Leu Val Phe Gln Ser  
 52 85 90 95  
 55 Ile Met Asp Ser Gly Gly Gly Leu Gly Gly Leu Pro Ser Asn Leu Gln  
 56 100 105 110  
 59 Thr Leu Phe Ser Gln Ala Tyr Ser Ala Gly Ala Arg Ile His Thr Asn  
 60 115 120 125  
 63 Ser Trp Gly Ala Ala Val Asn Gly Ala Tyr Thr Thr Asp Ser Arg Asn  
 64 130 135 140  
 67 Val Asp Asp Tyr Val Arg Lys Asn Asp Met Thr Ile Leu Phe Ala Ala  
 68 145 150 155 160  
 71 Gly Asn Glu Gly Pro Asn Gly Gly Thr Ile Ser Ala Pro Gly Thr Ala  
 72 165 170 175  
 75 Lys Asn Ala Ile Thr Val Gly Ala Thr Glu Asn Leu Arg Pro Ser Phe  
 76 180 185 190

ENTERED

## RAW SEQUENCE LISTING

DATE: 10/06/2004

PATENT APPLICATION: US/10/820,714A

TIME: 12:22:57

Input Set : A:\251697.ST25.txt

Output Set: N:\CRF4\10062004\J820714A.raw

```

79 Gly Ser Tyr Ala Asp Asn Ile Asn His Val Ala Gln Phe Ser Ser Arg
80      195      200      205
83 Gly Pro Thr Lys Asp Gly Arg Ile Lys Pro Asp Val Met Ala Pro Gly
84      210      215      220
87 Thr Phe Ile Leu Ser Ala Arg Ser Ser Leu Ala Pro Asp Ser Ser Phe
88 225      230      235      240
91 Trp Ala Asn His Asp Ser Lys Tyr Ala Tyr Met Gly Gly Thr Ser Met
92      245      250      255
95 Ala Thr Pro Ile Val Ala Gly Asn Val Ala Gln Leu Arg Glu His Phe
96      260      265      270
99 Val Lys Asn Arg Gly Ile Thr Pro Lys Pro Ser Leu Leu Lys Ala Ala
100     275      280      285
103 Leu Ile Ala Gly Ala Ala Asp Ile Gly Leu Gly Tyr Pro Asn Gly Asn
104     290      295      300
107 Gln Gly Trp Gly Arg Val Thr Leu Asp Lys Ser Leu Asn Val Ala Tyr
108 305      310      315      320
111 Val Asn Glu Ser Ser Ser Leu Ser Thr Ser Gln Lys Ala Thr Tyr Ser
112      325      330      335
115 Phe Thr Ala Thr Ala Gly Lys Pro Leu Lys Ile Ser Leu Val Trp Ser
116      340      345      350
119 Asp Ala Pro Ala Ser Thr Thr Ala Ser Val Thr Leu Val Asn Asp Leu
120      355      360      365
123 Asp Leu Val Ile Thr Ala Pro Asn Gly Thr Gln Tyr Val Gly Asn Asp
124      370      375      380
127 Phe Thr Ser Pro Tyr Asn Asp Asn Trp Asp Gly Arg Asn Asn Val Glu
128 385      390      395      400
131 Asn Val Phe Ile Asn Ala Pro Gln Ser Gly Thr Tyr Thr Ile Glu Val
132      405      410      415
135 Gln Ala Tyr Asn Val Pro Val Gly Pro Gln Thr Phe Ser Leu Ala Ile
136      420      425      430
139 Val Asn
143 <210> SEQ ID NO: 2
144 <211> LENGTH: 1923
145 <212> TYPE: DNA
146 <213> ORGANISM: Bacillus sp. KSM-KP43
149 <220> FEATURE:
150 <221> NAME/KEY: CDS
151 <222> LOCATION: (1)..(1920)
153 <220> FEATURE:
154 <221> NAME/KEY: sig_peptide
155 <222> LOCATION: (1)..(618)
157 <220> FEATURE:
158 <221> NAME/KEY: mat_peptide
159 <222> LOCATION: (619)..()
161 <400> SEQUENCE: 2
162 atg aga aag aag aaa aag gtg ttt tta tct gtt tta tca gct gca      45
163 Met Arg Lys Lys Lys Lys Val Phe Leu Ser Val Leu Ser Ala Ala
164 -205 -200 -195
166 gcg att ttg tcg act gtt gcg tta agt aat cca tct gca ggt ggt      90

```

## RAW SEQUENCE LISTING

DATE: 10/06/2004

PATENT APPLICATION: US/10/820,714A

TIME: 12:22:57

Input Set : A:\251697.ST25.txt

Output Set: N:\CRF4\10062004\J820714A.raw

167	Ala	Ile	Leu	Ser	Thr	Val	Ala	Leu	Ser	Asn	Pro	Ser	Ala	Gly	Gly		
168		-190					-185					-180					
170	gca	agg	aat	ttt	gat	ctg	gat	ttc	aaa	gga	att	cag	aca	aca	act	135	
171	Ala	Arg	Asn	Phe	Asp	Leu	Asp	Phe	Lys	Gly	Ile	Gln	Thr	Thr	Thr		
172		-175					-170					-165					
174	gat	gct	aaa	ggt	ttc	tcc	aag	cag	ggg	cag	act	ggt	gct	gct	gct	180	
175	Asp	Ala	Lys	Gly	Phe	Ser	Lys	Gln	Gly	Gln	Thr	Gly	Ala	Ala	Ala		
176		-160					-155					-150					
178	ttt	ctg	gtg	gaa	tct	gaa	aat	gtg	aaa	ctc	cca	aaa	ggt	ttg	cag	225	
179	Phe	Leu	Val	Glu	Ser	Glu	Asn	Val	Lys	Leu	Pro	Lys	Gly	Leu	Gln		
180		-145					-140					-135					
182	aag	aag	ctt	gaa	aca	gtc	ccg	gca	aat	aat	aaa	ctc	cat	att	atc	270	
183	Lys	Lys	Leu	Glu	Thr	Val	Pro	Ala	Asn	Asn	Lys	Leu	His	Ile	Ile		
184		-130					-125					-120					
186	caa	ttc	aat	gga	cca	att	tta	gaa	gaa	aca	aaa	cag	cag	ctg	gaa	315	
187	Gln	Phe	Asn	Gly	Pro	Ile	Leu	Glu	Glu	Thr	Lys	Gln	Gln	Leu	Glu		
188		-115					-110					-105					
190	aaa	aca	ggg	gca	aag	att	ctc	gac	tac	ata	cct	gat	tat	gct	tac	att	363
191	Lys	Thr	Gly	Ala	Lys	Ile	Leu	Asp	Tyr	Ile	Pro	Asp	Tyr	Ala	Tyr	Ile	
192		-100					-95					-90					
194	gtc	gag	tat	gag	ggc	gat	gtt	aag	tca	gca	aca	agc	acc	att	gag	cac	411
195	Val	Glu	Tyr	Glu	Gly	Asp	Val	Lys	Ser	Ala	Thr	Ser	Thr	Ile	Glu	His	
196		-85				-80				-75			-70				
198	gtg	gaa	tcc	gtg	gag	cct	tat	ttg	ccg	ata	tac	aga	ata	gat	ccc	cag	459
199	Val	Glu	Ser	Val	Glu	Pro	Tyr	Leu	Pro	Ile	Tyr	Arg	Ile	Asp	Pro	Gln	
200			-65				-60					-55					
202	ctt	ttc	aca	aaa	ggg	gca	tca	gag	ctt	gta	aaa	gca	gtg	gcg	ctt	gat	507
203	Leu	Phe	Thr	Lys	Gly	Ala	Ser	Glu	Leu	Val	Lys	Ala	Val	Ala	Leu	Asp	
204		-50					-45					-40					
206	aca	aag	cag	aaa	aat	aaa	gag	gtg	caa	tta	aga	ggc	atc	gaa	caa	atc	555
207	Thr	Lys	Gln	Lys	Asn	Lys	Glu	Val	Gln	Leu	Arg	Gly	Ile	Glu	Gln	Ile	
208		-35					-30					-25					
210	gca	caa	ttc	gca	ata	agc	aat	gat	gtg	cta	tat	att	acg	gca	aag	cct	603
211	Ala	Gln	Phe	Ala	Ile	Ser	Asn	Asp	Val	Leu	Tyr	Ile	Thr	Ala	Lys	Pro	
212		-20					-15					-10					
214	gag	tat	aag	gtg	atg	aat	gat	gtt	gcg	cgt	gga	att	gtc	aaa	gcg	gat	651
215	Glu	Tyr	Lys	Val	Met	Asn	Asp	Val	Ala	Arg	Gly	Ile	Val	Lys	Ala	Asp	
216	-5			-1	1			5				10					
218	gtg	gct	cag	agc	agc	tac	ggg	ttg	tat	gga	caa	gga	cag	atc	gta	gcg	699
219	Val	Ala	Gln	Ser	Ser	Tyr	Gly	Leu	Tyr	Gly	Gln	Gly	Gln	Ile	Val	Ala	
220			15				20					25					
222	gtt	gcc	gat	aca	ggg	ctt	gat	aca	ggt	cgc	aat	gac	agt	tcg	atg	cat	747
223	Val	Ala	Asp	Thr	Gly	Leu	Asp	Thr	Gly	Arg	Asn	Asp	Ser	Ser	Met	His	
224		30					35					40					
226	gaa	gcc	ttc	cgc	ggg	aaa	att	act	gca	tta	tat	gca	ttg	gga	cgg	acg	795
227	Glu	Ala	Phe	Arg	Gly	Lys	Ile	Thr	Ala	Leu	Tyr	Ala	Leu	Gly	Arg	Thr	
228		45					50					55					
230	aat	aat	gcc	aat	gat	acg	aat	ggt	cat	ggt	acg	cat	gtg	gct	ggc	tcc	843
231	Asn	Asn	Ala	Asn	Asp	Thr	Asn	Gly	His	Gly	Thr	His	Val	Ala	Gly	Ser	

## RAW SEQUENCE LISTING

DATE: 10/06/2004

PATENT APPLICATION: US/10/820,714A

TIME: 12:22:57

Input Set : A:\251697.ST25.txt

Output Set: N:\CRF4\10062004\J820714A.raw

232	60				65				70				75				
234	gta	tta	gga	aac	ggc	tcc	act	aat	aaa	gga	atg	gcg	cct	cag	gcg	aat	891
235	Val	Leu	Gly	Asn	Gly	Ser	Thr	Asn	Lys	Gly	Met	Ala	Pro	Gln	Ala	Asn	
236				80					85				90				
238	cta	gtc	ttc	caa	tct	atc	atg	gat	agc	ggg	gga	ctt	gga	gga	cta		939
239	Leu	Val	Phe	Gln	Ser	Ile	Met	Asp	Ser	Gly	Gly	Gly	Leu	Gly	Gly	Leu	
240				95					100				105				
242	cct	tcg	aat	ctg	caa	acc	tta	ttc	agc	caa	gca	tac	agt	gct	ggg	gcc	987
243	Pro	Ser	Asn	Leu	Gln	Thr	Leu	Phe	Ser	Gln	Ala	Tyr	Ser	Ala	Gly	Ala	
244			110					115				120					
246	aga	att	cat	aca	aac	tcc	tgg	gga	gca	gca	gtg	aat	ggg	gct	tac	aca	1035
247	Arg	Ile	His	Thr	Asn	Ser	Trp	Gly	Ala	Ala	Val	Asn	Gly	Ala	Tyr	Thr	
248		125					130				135						
250	aca	gat	tcc	aga	aat	gtg	gat	gac	tat	gtg	cgc	aaa	aat	gat	atg	acg	1083
251	Thr	Asp	Ser	Arg	Asn	Val	Asp	Asp	Tyr	Val	Arg	Lys	Asn	Asp	Met	Thr	
252	140				145				150				155				
254	atc	ctt	ttc	gct	gcc	ggg	aat	gaa	gga	ccg	aac	ggc	gga	acc	atc	agt	1131
255	Ile	Leu	Phe	Ala	Ala	Gly	Asn	Glu	Gly	Pro	Asn	Gly	Gly	Thr	Ile	Ser	
256				160					165				170				
258	gca	cca	ggc	aca	gct	aaa	aat	gca	ata	aca	gtc	gga	gct	acg	gaa	aac	1179
259	Ala	Pro	Gly	Thr	Ala	Lys	Asn	Ala	Ile	Thr	Val	Gly	Ala	Thr	Glu	Asn	
260			175					180				185					
262	ctc	cgc	cca	agc	ttt	ggg	tct	tat	gcg	gac	aat	atc	aac	cat	gtg	gca	1227
263	Leu	Arg	Pro	Ser	Phe	Gly	Ser	Tyr	Ala	Asp	Asn	Ile	Asn	His	Val	Ala	
264		190					195				200						
266	cag	ttc	tct	tca	cgt	gga	ccg	aca	aag	gat	gga	cgg	atc	aaa	ccg	gat	1275
267	Gln	Phe	Ser	Ser	Arg	Gly	Pro	Thr	Lys	Asp	Gly	Arg	Ile	Lys	Pro	Asp	
268		205				210				215							
270	gtc	atg	gca	ccg	gga	acg	ttc	ata	cta	tca	gca	aga	tct	tct	ctt	gca	1323
271	Val	Met	Ala	Pro	Gly	Thr	Phe	Ile	Leu	Ser	Ala	Arg	Ser	Ser	Leu	Ala	
272	220				225				230				235				
274	ccg	gat	tcc	tcc	ttc	tgg	gcg	aac	cat	gac	agt	aaa	tat	gca	tac	atg	1371
275	Pro	Asp	Ser	Ser	Phe	Trp	Ala	Asn	His	Asp	Ser	Lys	Tyr	Ala	Tyr	Met	
276			240					245			250						
278	ggg	gga	acg	tcc	atg	gct	aca	ccg	atc	gtt	gct	gga	aac	gtg	gca	cag	1419
279	Gly	Gly	Thr	Ser	Met	Ala	Thr	Pro	Ile	Val	Ala	Gly	Asn	Val	Ala	Gln	
280			255					260			265						
282	ctt	cgt	gag	cat	ttt	gtg	aaa	aac	aga	ggc	atc	aca	cca	aag	cct	tct	1467
283	Leu	Arg	Glu	His	Phe	Val	Lys	Asn	Arg	Gly	Ile	Thr	Pro	Lys	Pro	Ser	
284		270					275				280						
286	cta	tta	aaa	gcg	gca	ctg	att	gcc	ggg	gca	gct	gac	atc	ggc	ctt	ggc	1515
287	Leu	Leu	Lys	Ala	Ala	Leu	Ile	Ala	Gly	Ala	Ala	Asp	Ile	Gly	Leu	Gly	
288		285				290			295								
290	tac	ccg	aac	ggg	aac	caa	gga	tgg	gga	cga	gtg	aca	ttg	gat	aaa	tcc	1563
291	Tyr	Pro	Asn	Gly	Asn	Gln	Gly	Trp	Gly	Arg	Val	Thr	Leu	Asp	Lys	Ser	
292	300				305				310				315				
294	ctg	aac	gtt	gcc	tat	gtg	aac	gag	tcc	agt	tct	cta	tcc	acc	agc	caa	1611
295	Leu	Asn	Val	Ala	Tyr	Val	Asn	Glu	Ser	Ser	Ser	Leu	Ser	Thr	Ser	Gln	
296				320				325				330					

## RAW SEQUENCE LISTING

DATE: 10/06/2004

PATENT APPLICATION: US/10/820,714A

TIME: 12:22:57

Input Set : A:\251697.ST25.txt

Output Set: N:\CRF4\10062004\J820714A.raw

```

298 aaa gcg acg tac tcg ttt act gct act gcc ggc aag cct ttg aaa atc      1659
299 Lys Ala Thr Tyr Ser Phe Thr Ala Thr Ala Gly Lys Pro Leu Lys Ile
300          335          340          345
302 tcc ctg gta tgg tct gat gcc cct gcg agc aca act gct tcc gta acg      1707
303 Ser Leu Val Trp Ser Asp Ala Pro Ala Ser Thr Thr Ala Ser Val Thr
304          350          355          360
306 ctt gtc aat gat ctg gac ctt gtc att acc gct cca aat ggc aca cag      1755
307 Leu Val Asn Asp Leu Asp Leu Val Ile Thr Ala Pro Asn Gly Thr Gln
308          365          370          375
310 tat gta gga aat gac ttt act tcg cca tac aat gat aac tgg gat ggc      1803
311 Tyr Val Gly Asn Asp Phe Thr Ser Pro Tyr Asn Asp Asn Trp Asp Gly
312 380          385          390          395
314 cgc aat aac gta gaa aat gta ttt att aat gca cca caa agc ggg acg      1851
315 Arg Asn Asn Val Glu Asn Val Phe Ile Asn Ala Pro Gln Ser Gly Thr
316          400          405          410
318 tat aca att gag gta cag gct tat aac gta ccg gtt gga cca cag acc      1899
319 Tyr Thr Ile Glu Val Gln Ala Tyr Asn Val Pro Val Gly Pro Gln Thr
320          415          420          425
322 ttc tcg ttg gca att gtg aat taa      1923
323 Phe Ser Leu Ala Ile Val Asn
324          430
327 <210> SEQ ID NO: 3
328 <211> LENGTH: 640
329 <212> TYPE: PRT
330 <213> ORGANISM: Bacillus sp. KSM-KP43
332 <400> SEQUENCE: 3
334 Met Arg Lys Lys Lys Lys Val Phe Leu Ser Val Leu Ser Ala Ala
335 -205          -200          -195
338 Ala Ile Leu Ser Thr Val Ala Leu Ser Asn Pro Ser Ala Gly Gly
339 -190          -185          -180
342 Ala Arg Asn Phe Asp Leu Asp Phe Lys Gly Ile Gln Thr Thr Thr
343 -175          -170          -165
346 Asp Ala Lys Gly Phe Ser Lys Gln Gly Gln Thr Gly Ala Ala Ala
347 -160          -155          -150
350 Phe Leu Val Glu Ser Glu Asn Val Lys Leu Pro Lys Gly Leu Gln
351 -145          -140          -135
354 Lys Lys Leu Glu Thr Val Pro Ala Asn Asn Lys Leu His Ile Ile
355 -130          -125          -120
358 Gln Phe Asn Gly Pro Ile Leu Glu Glu Thr Lys Gln Gln Leu Glu
359 -115          -110          -105
362 Lys Thr Gly Ala Lys Ile Leu Asp Tyr Ile Pro Asp Tyr Ala Tyr Ile
363 -100          -95          -90
366 Val Glu Tyr Glu Gly Asp Val Lys Ser Ala Thr Ser Thr Ile Glu His
367 -85          -80          -75          -70
370 Val Glu Ser Val Glu Pro Tyr Leu Pro Ile Tyr Arg Ile Asp Pro Gln
371          -65          -60          -55
374 Leu Phe Thr Lys Gly Ala Ser Glu Leu Val Lys Ala Val Ala Leu Asp
375          -50          -45          -40
378 Thr Lys Gln Lys Asn Lys Glu Val Gln Leu Arg Gly Ile Glu Gln Ile

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/820,714A

DATE: 10/06/2004

TIME: 12:22:58

Input Set : A:\251697.ST25.txt

Output Set: N:\CRF4\10062004\J820714A.raw